

5/17

#11

ENTERED



OIPE

## RAW SEQUENCE LISTING

DATE: 05/22/2002

PATENT APPLICATION: US/09/900,345A

TIME: 11:27:26

Input Set : A:\10338\_5US.txt

Output Set: N:\CRF3\05222002\I900345A.raw

```

3 <110> APPLICANT: Frazer, Ian Hector
4   Zhou, Jian
6 <120> TITLE OF INVENTION: METHOD AND POLYNUCLEOTIDES FOR DETERMINING TRANSLATIONAL
7   EFFICIENCY OF A CODON
9 <130> FILE REFERENCE: 10338-5US
11 <140> CURRENT APPLICATION NUMBER: 09/900,345A
12 <141> CURRENT FILING DATE: 2001-07-06
14 <150> PRIOR APPLICATION NUMBER: AU PP8078
15 <151> PRIOR FILING DATE: 1999-01-08
17 <150> PRIOR APPLICATION NUMBER: PCT/AU00/00008
18 <151> PRIOR FILING DATE: 2000-01-07
20 <160> NUMBER OF SEQ ID NOS: 185
22 <170> SOFTWARE: PatentIn Ver. 2.0
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 732
26 <212> TYPE: DNA
27 <213> ORGANISM: Artificial Sequence
29 <220> FEATURE:
30 <223> OTHER INFORMATION: Description of Artificial Sequence: Ala(GCA)5GFP
32 <220> FEATURE:
33 <221> NAME/KEY: CDS
34 <222> LOCATION: (1)..(732)
36 <400> SEQUENCE: 1
37 atg gca gca gca gca gca agc aag ggc gag gaa ctg ttc act ggc gtg   48
38 Met Ala Ala Ala Ala Ala Ser Lys Gly Glu Glu Leu Phe Thr Gly Val
39   1           5           10           15
41 gtc cca att ctc gtg gaa ctg gat ggc gat gtg aat ggg cac aaa ttt   96
42 Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe
43   20           25           30
45 tct gtc agc gga gag ggt gaa ggt gat gcc aca tac gga aag ctc acc   144
46 Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr
47   35           40           45
49 ctg aaa ttc atc tgc acc act gga aag ctc cct gtg cca tgg cca aca   192
50 Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr
51   50           55           60
53 ctg gtc act acc ttc tct tat ggc gtg cag tgc ttt tcc aga tac cca   240
54 Leu Val Thr Thr Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro
55   65           70           75           80
57 gac cat atg aag cag cat gac ttt ttc aag agc gcc atg ccc gag ggc   288
58 Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly
59   85           90           95
61 tat gtg cag gag aga acc atc ttt ttc aaa gat gac ggg aac tac aag   336
62 Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys

```

## RAW SEQUENCE LISTING

DATE: 05/22/2002

PATENT APPLICATION: US/09/900,345A

TIME: 11:27:26

Input Set : A:\10338\_5US.txt

Output Set: N:\CRF3\05222002\I900345A.raw

```

63          100          105          110
65 acc cgc gct gaa gtc aag ttc gaa ggt gac acc ctg gtg aat aga atc 384
66 Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile
67          115          120          125
69 gag ctg aag ggc att gac ttt aag gag gat gga aac att ctc ggc cac 432
70 Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His
71          130          135          140
73 aag ctg gaa tac aac tat aac tcc cac aat gtg tac atc atg gcc gac 480
74 Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp
75 145          150          155          160
77 aag caa aag aat ggc atc aag gtc aac ttc aag atc aga cac aac att 528
78 Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile
79          165          170          175
81 gag gat gga tcc gtg cag ctg gcc gac cat tat caa cag aac act cca 576
82 Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro
83          180          185          190
85 atc ggc gac ggc cct gtg ctc ctc cca gac aac cat tac ctg tcc acc 624
86 Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr
87          195          200          205
89 cag tct gcc ctg tct aaa gat ccc aac gaa aag aga gac cac atg gtc 672
90 Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val
91          210          215          220
93 ctg ctg gag ttt gtg acc gct gct ggg atc aca cat ggc atg gac gag 720
94 Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu
95 225          230          235          240
97 ctg tac aag tga 732
98 Leu Tyr Lys
101 <210> SEQ ID NO: 2
102 <211> LENGTH: 243
103 <212> TYPE: PRT
104 <213> ORGANISM: Artificial Sequence
106 <220> FEATURE:
107 <223> OTHER INFORMATION: Description of Artificial Sequence: Ala(GCA)5GFP
109 <400> SEQUENCE: 2
110 Met Ala Ala Ala Ala Ala Ser Lys Gly Glu Glu Leu Phe Thr Gly Val
111 1          5          10          15
113 Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe
114          20          25          30
116 Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr
117          35          40          45
119 Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr
120          50          55          60
122 Leu Val Thr Thr Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro
123 65          70          75          80
125 Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly
126          85          90          95
128 Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys
129          100          105          110
131 Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile

```

## RAW SEQUENCE LISTING

DATE: 05/22/2002

PATENT APPLICATION: US/09/900,345A

TIME: 11:27:26

Input Set : A:\10338\_5US.txt

Output Set: N:\CRF3\05222002\I900345A.raw

```

132          115          120          125
134 Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His
135          130          135          140
137 Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp
138 145          150          155          160
140 Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile
141          165          170          175
143 Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro
144          180          185          190
146 Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr
147          195          200          205
149 Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val
150          210          215          220
152 Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu
153 225          230          235          240
155 Leu Tyr Lys
158 <210> SEQ ID NO: 3
159 <211> LENGTH: 732
160 <212> TYPE: DNA
161 <213> ORGANISM: Artificial Sequence
163 <220> FEATURE:
164 <223> OTHER INFORMATION: Description of Artificial Sequence: Ala(GCC)5GFP
166 <220> FEATURE:
167 <221> NAME/KEY: CDS
168 <222> LOCATION: (1)..(732)
170 <400> SEQUENCE: 3
171 atg gcc gcc gcc gcc gcc agc aag ggc gag gaa ctg ttc act ggc gtg 48
172 Met Ala Ala Ala Ala Ala Ser Lys Gly Glu Glu Leu Phe Thr Gly Val
173 1 5 10 15
175 gtc cca att ctc gtg gaa ctg gat ggc gat gtg aat ggg cac aaa ttt 96
176 Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe
177 20 25 30
179 tct gtc agc gga gag ggt gaa ggt gat gcc aca tac gga aag ctc acc 144
180 Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr
181 35 40 45
183 ctg aaa ttc atc tgc acc act gga aag ctc cct gtg cca tgg cca aca 192
184 Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr
185 50 55 60
187 ctg gtc act acc ttc tct tat ggc gtg cag tgc ttt tcc aga tac cca 240
188 Leu Val Thr Thr Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro
189 65 70 75 80
191 gac cat atg aag cag cat gac ttt ttc aag agc gcc atg ccc gag ggc 288
192 Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly
193 85 90 95
195 tat gtg cag gag aga acc atc ttt ttc aaa gat gac ggg aac tac aag 336
196 Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys
197 100 105 110
199 acc cgc gct gaa gtc aag ttc gaa ggt gac acc ctg gtg aat aga atc 384
200 Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile

```

## RAW SEQUENCE LISTING.

DATE: 05/22/2002

PATENT APPLICATION: US/09/900,345A

TIME: 11:27:26

Input Set : A:\10338\_5US.txt

Output Set: N:\CRF3\05222002\I900345A.raw

```

201          115          120          125
203 gag ctg aag ggc att gac ttt aag gag gat gga aac att ctc ggc cac 432
204 Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His
205          130          135          140
207 aag ctg gaa tac aac tat aac tcc cac aat gtg tac atc atg gcc gac 480
208 Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp
209 145          150          155          160
211 aag caa aag aat ggc atc aag gtc aac ttc aag atc aga cac aac att 528
212 Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile
213          165          170          175
215 gag gat gga tcc gtg cag ctg gcc gac cat tat caa cag aac act cca 576
216 Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro
217          180          185          190
219 atc ggc gac ggc cct gtg ctc ctc cca gac aac cat tac ctg tcc acc 624
220 Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr
221          195          200          205
223 cag tct gcc ctg tct aaa gat ccc aac gaa aag aga gac cac atg gtc 672
224 Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val
225          210          215          220
227 ctg ctg gag ttt gtg acc gct gct ggg atc aca cat ggc atg gac gag 720
228 Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu
229 225          230          235          240
231 ctg tac aag tga 732
232 Leu Tyr Lys
235 <210> SEQ ID NO: 4
236 <211> LENGTH: 243
237 <212> TYPE: PRT
238 <213> ORGANISM: Artificial Sequence
240 <220> FEATURE:
241 <223> OTHER INFORMATION: Description of Artificial Sequence: Ala(GCC)5GFP
243 <400> SEQUENCE: 4
244 Met Ala Ala Ala Ala Ala Ser Lys Gly Glu Glu Leu Phe Thr Gly Val
245 1          5          10          15
247 Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe
248          20          25          30
250 Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr
251          35          40          45
253 Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr
254          50          55          60
256 Leu Val Thr Thr Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro
257 65          70          75          80
259 Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly
260          85          90          95
262 Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys
263          100          105          110
265 Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile
266          115          120          125
268 Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His
269          130          135          140

```

## RAW SEQUENCE LISTING

DATE: 05/22/2002

PATENT APPLICATION: US/09/900,345A

TIME: 11:27:26

Input Set : A:\10338\_5US.txt

Output Set: N:\CRF3\05222002\I900345A.raw

```

271 Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp
272 145 150 155 160
274 Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile
275 165 170 175
277 Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro
278 180 185 190
280 Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr
281 195 200 205
283 Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val
284 210 215 220
286 Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu
287 225 230 235 240
289 Leu Tyr Lys
292 <210> SEQ ID NO: 5
293 <211> LENGTH: 732
294 <212> TYPE: DNA
295 <213> ORGANISM: Artificial Sequence
297 <220> FEATURE:
298 <223> OTHER INFORMATION: Description of Artificial Sequence: Ala(GCG)5GFP
300 <220> FEATURE:
301 <221> NAME/KEY: CDS
302 <222> LOCATION: (1)..(732)
304 <400> SEQUENCE: 5
305 atg gcg gcg gcg gcg agc aag ggc gag gaa ctg ttc act ggc gtg 48
306 Met Ala Ala Ala Ala Ser Lys Gly Glu Glu Leu Phe Thr Gly Val
307 1 5 10 15
309 gtc cca att ctc gtg gaa ctg gat ggc gat gtg aat ggg cac aaa ttt 96
310 Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe
311 20 25 30
313 tct gtc agc gga gag ggt gaa ggt gat gcc aca tac gga aag ctc acc 144
314 Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr
315 35 40 45
317 ctg aaa ttc atc tgc acc act gga aag ctc cct gtg cca tgg cca aca 192
318 Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr
319 50 55 60
321 ctg gtc act acc ttc tct tat ggc gtg cag tgc ttt tcc aga tac cca 240
322 Leu Val Thr Thr Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro
323 65 70 75 80
325 gac cat atg aag cag cat gac ttt ttc aag agc gcc atg ccc gag ggc 288
326 Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly
327 85 90 95
329 tat gtg cag gag aga acc atc ttt ttc aaa gat gac ggg aac tac aag 336
330 Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys
331 100 105 110
333 acc cgc gct gaa gtc aag ttc gaa ggt gac acc ctg gtg aat aga atc 384
334 Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile
335 115 120 125
337 gag ctg aag ggc att gac ttt aag gag gat gga aac att ctc ggc cac 432
338 Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/900,345A

DATE: 05/22/2002  
TIME: 11:27:27

Input Set : A:\10338\_5US.txt  
Output Set: N:\CRF3\05222002\I900345A.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:2; Line(s) 107

**VERIFICATION SUMMARY**

DATE: 05/22/2002

PATENT APPLICATION: US/09/900,345A

TIME: 11:27:27

Input Set : A:\10338\_5US.txt

Output Set: N:\CRF3\05222002\I900345A.raw